Modelling the influence of biotic factors on species distribution patterns

Katie Leach [a,b]*, W. Ian Montgomery [a,b,c] and Neil Reid [b,c]

[a] *Quercus*, School of Biological Sciences, Queen's University Belfast, Belfast, BT9 7BL. Northern Ireland.

[b] School of Biological Sciences, Queen's University Belfast, Belfast, BT9 7BL. Northern Ireland.

[c] Institute for Global Food Security (IGFS), Queen's University Belfast, Belfast, BT9 5BN. Northern Ireland.

*Katie Leach: kleach01@qub.ac.uk

W. Ian Montgomery: i.montgomery@qub.ac.uk

Neil Reid: neil.reid@qub.ac.uk

* Corresponding Author

ABSTRACT

Biotic interactions can have large effects on species distributions yet their role in shaping species ranges is seldom explored due to historical difficulties in incorporating biotic factors into models without *a priori* knowledge on interspecific interactions. Improved SDMs, which account for biotic factors and do not require *a priori* knowledge on species interactions, are needed to fully understand species distributions. Here, we model the influence of abiotic and biotic factors on species distribution patterns and explore the robustness of distributions under future climate change. We fit hierarchical spatial models using Integrated Nested Laplace Approximation (INLA) for lagomorph species throughout Europe and test the predictive ability of models containing only abiotic factors against models containing abiotic and biotic factors. We account for residual spatial autocorrelation using a conditional autoregressive (CAR) model. Model outputs are used to estimate areas in which abiotic and biotic factors determine species' ranges. INLA models containing both abiotic and biotic factors had substantially better predictive ability than models containing abiotic factors only, for all but one of the four species. In models containing abiotic and biotic factors, both appeared equally important as determinants of lagomorph ranges, but the influences were spatially heterogeneous. Parts of widespread lagomorph ranges highly influenced by biotic factors will be less robust to future changes in climate, whereas parts of more localised species ranges highly influenced by the environment may be less robust to future climate. SDMs that do not explicitly include biotic factors are potentially misleading and omit a very important source of variation. For the field of species distribution modelling to advance, biotic factors must be taken into account in order to improve the reliability of predicting species distribution patterns both presently and under future climate change.

Keywords: *biotic interactions; climate; competition; INLA; Order Lagomorpha; species distribution modelling.*

1.1 INTRODUCTION

The factors shaping species spatial distributions are crucial for our understanding of patterns of biodiversity, and, hence, are commonly studied. Species distributions are constrained by abiotic conditions, for example, suitable climate, and biotic conditions, for example, competitive interspecific interactions, as well as movement behaviour, such as dispersal or colonisation (Soberon & Nakamura, 2009). Species Distribution Models (SDMs) are widely used in ecology and typically relate species occurrences to (abiotic) environmental variables to produce models of environmental suitability, which can be spatially or temporally extrapolated to unsurveyed areas or into past or future conditions (e.g. Elith *et al*., 2006). Most SDMs use only environmental variables to predict species distributions and often do not offer credible statistical tests of the relative influence of these abiotic factors. Typically, they also ignore other key determinants of species ranges, most pertinently, biotic interactions, as including these has tended to require *a priori* knowledge on species interactions (see Elith & Leathwick, 2009; Beale *et al*., 2014).

Although macroclimate drives distribution patterns at large scales (Whittaker, 1975; Pearson & Dawson, 2003; Thuiller *et al*., 2003), there is growing evidence of the important role of biotic interactions in delineating species' range extents (Jablonski, 2008; Wiens, 2011; Wisz *et al*., 2012). Biotic interactions can take place at local scales, for example, predation, parasitism, competition and disturbance, or regional scales, for example, dispersal, speciation, extinction and expansions or contractions of species ranges (Cornell & Lawton, 1992; Amarasekare, 2003). Competitive interactions, in particular, tend to produce biogeographical patterns in species distributions; species may meet at a sharp boundary with little or no overlap (Flux, 2008). Parapatric distributions, whereby two species have separate, but contiguous ranges, with no physical barrier between them and only co-occur, if at all, in a narrow contact zone, may result from such competitive interactions (Bull, 1991, Gutiérrez *et al*., 2014). In

63 order to fully understand species distributions, we need to recognise the role of biotic
64 interactions in shaping geographic ranges (Wisz *et al*., 2012) and develop novel SDM
65 techniques to disaggregate the effects of the abiotic environment and biotic interactions.

66 Despite the many limitations with current correlative SDM methods, they remain one of the
67 only methods to predict species distributions. Further, they have practical applications in
68 conservation management as well as providing insights into the past and future distributions of
69 organisms and the factors that shape their biogeography. However, if predictions do not
70 explicitly include biotic factors, they could potentially misinform conservation management
71 decisions. Accurate representations of species distributions are vital for the design and
72 implementation of appropriate conservation measures, e.g. protected areas, but SDMs are
73 unlikely to produce reliable predictions if they rely on abiotic factors alone (Hof *et al*., 2012;
74 Kissling *et al*., 2012).

75 Studies using SDMs are beginning to consider biotic interactions by including the
76 occurrence of other species' as additional covariates (e.g. Pellissier *et al*., 2010) and this has
77 been central in improving performance (Araujo & Luoto, 2007; Mod *et al*., 2015). Surrogates
78 for biotic-interaction gradients are also used; for example, habitat productivity, which is known
79 to be a basis of competition in plants (Maestre *et al*., 2010), or species richness patterns. SDMs
80 may also be hybridised with dynamic models, such as BIOMOVE (Midgley *et al*., 2010), but
81 these methods all require *a priori* ecological knowledge and can only be applied to one
82 interacting species' pair at a time (Wisz *et al*., 2012).

83 Markov Chain Monte Carlo (MCMC) simulations offer a flexible framework for species
84 distribution modelling (Beale *et al*., 2014) but they come with a wide range of problems
85 regarding convergence, processing time and implementation (Beguin *et al*., 2012). An
86 alternative for fitting such Bayesian hierarchical spatial models uses Integrated Nested Laplace

87 Approximation (INLA; Rue *et al*., 2009) which does not require advanced programming skills,
88 yields comparable results to MCMC and is rapid (Beguin *et al*., 2012). Both methods can model
89 spatial associations within the data using conditional autoregressive (CAR) models, but
90 because INLA substitutes accurate, deterministic approximations to posterior marginal
91 distributions in place of long MCMC simulations, in the majority of occasions INLA requires
92 less processing time to model spatial effects (Beguin *et al*., 2012). However, INLA is not yet a
93 completely adequate replacement for MCMC because some models, for example observation
94 effort models (Illian, 2013), cannot be fitted with INLA but can with MCMC. More recently
95 'Joint SDMs' have been introduced, which simultaneously estimate the ranges of multiple
96 coexisting species producing mixtures of possible species assemblages (Pollock *et al*., 2014;
97 Harris, 2014). Joint SDMs are a substantial step forward, but most do not directly take into
98 account spatial autocorrelation. A major advantage of using INLA is the ability to account for
99 spatial autocorrelation with relative ease. SDM studies are especially sensitive to spatial
100 autocorrelation, which arises when nearby locations are not independent of each other
101 (Legendre, 1993), and if spatial autocorrelation is not considered it can result in misleading
102 and biased models (Lennon, 2000; Beale *et al*., 2010).

103 Extending the use of hierarchical spatial models to project species distributions under future
104 climate scenarios is problematic as species interactions can be unstable in space or time (Wisz
105 *et al*., 2012). Climatic changes may impact the distribution of interacting species which could,
106 in turn, alter interactions (Wisz *et al*., 2012), but species interactions may also influence the
107 impacts of climate change, for example by altering community dynamics (Gilman *et al*., 2010).
108 If interactions have minimal effects on predicted distributions then projection may be possible,
109 however, if interactions are influential then projection is difficult because the strength and/or
110 direction of these interactions is likely to change (Beale *et al*., 2014). Due to the difficulties
111 with quantitatively modelling the multiple impacts of future change on interacting species,

proxies for studying climate change impacts, i.e. using measures which indicate change, for example ecoregional climatic stability, can be useful in determining likely effects (Iwamura *et al*., 2010; Iwamura *et* al., 2013; Watson *et al*., 2013). Regions highly influenced by biotic interactions with a low robustness to future climate change (or low climatic stability) may be particularly susceptible to altered distributions and/ or interactions.

Here, we fit hierarchical spatial models using INLA to quantify the influence of biotic factors on species distribution patterns and test the difference in predictive ability between INLA models containing only abiotic factors and models containing abiotic and biotic factors. Using a number of explanatory factors, including climatic variables, co-occurring species, vegetation, topography and human influence, we aim to produce more reliable predictions of species distributions as well as estimating the areas in which abiotic and biotic factors determine species' ranges. *Post-hoc* tests exploring the climatic stability of these areas can highlight, for example, if areas influenced to a greater extent by biotic factors are likely to be less robust to future changes, and, therefore, direct future management actions towards either the conservation of co-occurring species or the conservation of particular habitats. This approach could be implemented widely, for a number of species, but here we use lagomorphs in Europe, a fairly large spatial extent, as a test dataset because their biotic interactions are well known (see Acevedo *et al*., 2012 and Leach *et al*., 2015a) and can, therefore, be easily validated. There are six lagomorph species found throughout Europe (*Lepus castroviejoi, L. corsicanus, L. europaeus, L. granatensis, L. timidus* and *Oryctolagus cuniculus*) and they occupy a wide range of environmental conditions. We expect large influences of biotic factors on lagomorph species distributions, and given the wide-ranging climatic conditions occupied, we anticipate significant implications for conservation management in a changing climate. We expect to recover competitive interactions between *L. europaeus* and *L. timidus*, given studies in Ireland and across Europe (Thulin, 2003; Reid & Montgomery, 2007; Reid, 2011; Caravaggi

137 *et al*., 2014; 2016), and also between *L. europaeus* and *L. granatensis*, given their parapatric
138 distributions and field studies in Iberia (Gortazar *et al*., 2007).

1.2 MATERIALS AND METHODS

1.2.1 Species data

International Union for Conservation of Nature (IUCN) geographic range polygons for European lagomorph species (Fig. 1) were rasterised in R version 3.1.1 at 30 arc-minute resolution (~50km^2 grid cells), with a value of 1 for species presence and 0 for absence. IUCN polygons have been used in a number of species distribution modelling studies to date (e.g. Lawler *et al*., 2009; Visconti *et al*., 2015), and whilst they may have higher omission errors (Graham & Hijmans, 2006; Murray *et al*., 2011), the detailed construction of the polygons together with the internal review process and expert assessments by the IUCN can lead to the production of more realistic distribution models. To illustrate the consequences of using different input data, species distributions models will be built with IUCN polygons and compared to those built using point occurrence data. For this exercise, point data was downloaded from the Global Biodiversity Information Facility (GBIF) Data Portal (http://data.gbif.org) and prepared as per Leach *et al*. (2015b). Models for two highly range-restricted species: *L. castroviejoi* and *L. corsicanus*, were extremely poor and are not considered further.

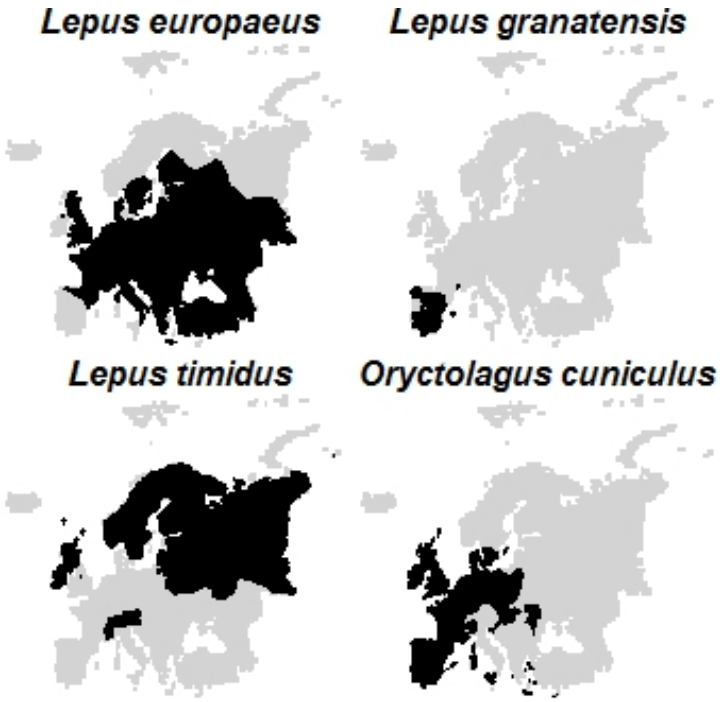


Fig. 1. IUCN geographic range polygons (black) for the European lagomorph species modelled. *L. castroviejoi* and *L. corsicanus* models were extremely poor and so their ranges are not shown here.

1.2.2 Environmental data

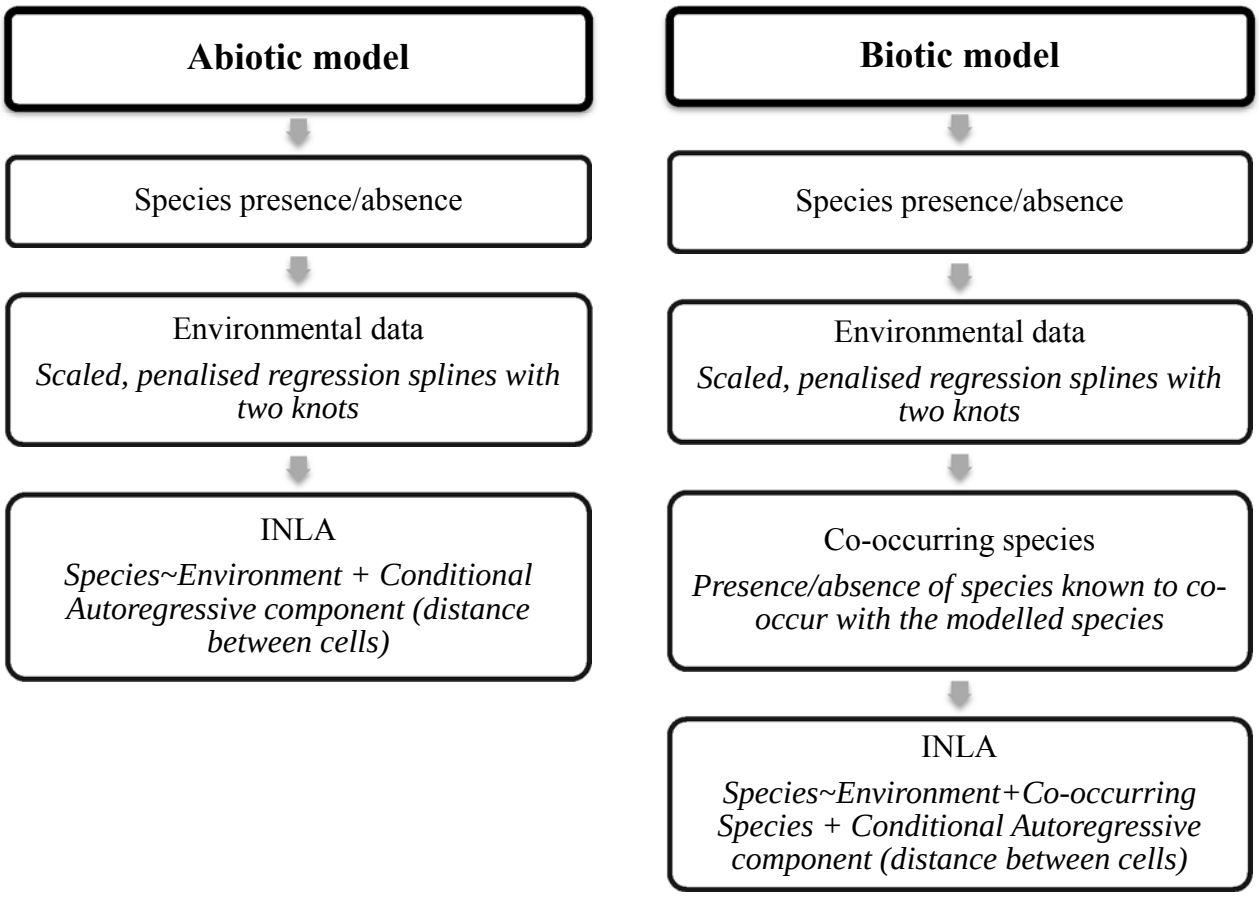
Current climate variables (~1950-2000) known to determine lagomorph species distributions were downloaded from WorldClim (www.worldclim.org) and resampled to the same resolution as the species data. Evapotranspiration was calculated using the Hargreaves equation (see Leach *et al*., 2015b for more details) and annual water balance was calculated by subtracting annual evapotranspiration from mean annual precipitation. The number of months with a Positive Water Balance (PWB) was calculated by subtracting each monthly evapotranspiration from its corresponding monthly precipitation, then converting into a binary format, where a value greater than zero was given a value of one and a value less than zero was kept at zero, finally summing the twelve binary scores (Kremen *et al*., 2008). Mean annual Normalised

Difference Vegetation Index (NDVI) was calculated from monthly values which were downloaded from the European Distributed Institute of Taxonomy (EDIT) Geoplatform (http://edit.csic.es/Soil-Vegetation-LandCover.html). Hilliness, an index of surface roughness, was calculated by finding the difference between maximum and minimum gradient values, based on a global Digital Elevation Model at 30 arc-minute resolution (Newton-Cross *et al*., 2007). Human Influence Index data were downloaded from the NASA Socioeconomic Data and Applications Centre (SEDAC) website (http://sedac.ciesin.columbia.edu/; WCS CIESIN, 2005). Subsequently, correlated environmental variables (minimum precipitation, minimum temperature, mean annual precipitation, mean annual temperature, solar radiation, annual water balance and annual evapotranspiration) were removed, leaving the following: maximum temperature, temperature seasonality, maximum precipitation, precipitation seasonality, PWB, NDVI, Hilliness and Human Influence Index. All environmental variables were rescaled to have a mean of zero and unit variance. Environmental data were extracted for each grid cell across Europe and combined in a dataset with species presence/absence data.

1.2.3 Model structure

Hierarchical spatial models were fit using INLA (Fig. 2). Code and sample data are available from: https://github.com/katieleach/BioticSDMs. Relationships between the response variable and individual covariates were restricted to functional forms with simple shapes (Austin, 2002) and the flexibility of these forms was part of the modelling approach. Each covariate was fitted as a smooth term represented by a penalised regression spline with two knots to describe biologically realistic unimodal response terms; regression splines were recombined to obtain fitted values for environmental variables as per Beale *et al*. (2014). Greater or lesser flexibility can be permitted by altering the number of knots in the spline terms. The spline models are detailed further in the R script. Species known to co-occur with the target (modelled) species

195 i.e. response variable were included as biotic covariates in addition to the environmental
196 variables. Residual spatial autocorrelation was accounted for using an intrinsic conditional
197 autoregressive (iCAR) error structure. This represents a spatial random effect and accounts for
198 spatially structured residuals, e.g. missing abiotic and biotic variables and processes, not
199 modelled by the included covariates. A vague Gaussian prior was assumed for the model
200 parameters and values of the shape parameters were defined so that the ratio of data points to
201 the number of effective parameters was >20; for more details see Beguin *et al*. (2012).



202

203 **Fig. 2.** Conceptual diagram of abiotic and biotic models.

1.2.4 Model evaluation

Predictive ability of species distribution models is typically assessed using metrics such as Area Under the Curve (AUC) or alternatives such as the True Skill Statistic (TSS) or Kappa values. However, these metrics have been criticised when using presence/pseudo-absence data and have limitations with respect to sensitivity, specificity, omission and commission errors (Allouche *et al*., 2006). These performance metrics require arbitrary designation of training and test data, and comparison of values between models for hypothesis testing is difficult. In Bayesian statistics, common methods of model evaluation include Bayesian Information Criterion (BIC; unavailable in R-INLA), Deviance Information Criterion (DIC) and Watanabe-Akaike information criterion (WAIC). But, all these methods over fit because of their equivalence to cross-validation (Plummer, 2008; Gelman *et al*., 2014). WAIC has recently been made available in R-INLA for model comparison and is recommended because it is a more fully Bayesian approach for out-of-sample prediction (Watanabe, 2010; Gelman *et al*., 2014), however, WAIC is not recommended for spatial analyses (Gelman *et al*., 2014). DIC is known to penalise for model complexity (Plummer, 2008), but Spiegelhalter *et al*. (2002) recommend DIC for spatial analyses. Given that there is no perfect solution to date, we evaluate models containing only abiotic factors against models containing both abiotic *and* biotic factors using both WAIC and DIC.

1.2.5 Analysis

All analyses were carried out in R version 3.1.1. Separate models were fit for each lagomorph species, with the abiotic *and* biotic model for one species defined as:

$$logit(P_i) = e_0 + (e_1 \times env1_i + e_2 \times env2_i) + (b_1 \times species1_i + \cdots b_2 \times species2_i) + SE_i$$

(*Equation* 1),

and the abiotic only model defined as:

$$logit(P_i) = e_0 + (e_1 \times env1_i + \cdots e_2 \times env2_i) + SE_i$$

(*Equation* 2),

where P_i was the probability of presence in cell i, e_0 was the intercept, e_1 and e_2 the estimated parameters for abiotic factors, $env1_i$and $env2_i$ the two spline bases for e_1 and e_2 in cell i, b_1 and b_2 the estimated parameters for biotic factors, $species1_i$ and $species2_i$ the two spline bases for b_1 and b_2 in cell i, and SE was the estimated spatially explicit error term for cell i. The total number of fixed-effect parameters differs between models due to differing numbers of co-occurring species; however, there were always 17 fixed-effect abiotic parameters (intercept plus two parameters for each abiotic factor – due to the two knots in the penalised regression spline).

Equations 1 and 2 can be separated into multiple terms in order to explore the relative influence of abiotic and biotic factors in each cell. Contributions of each term to the model were calculated as follows: the contribution of the spatially explicit error term or spatially structured residuals in cell i, $SE_{COMP_i} = inv.logit(SE_i)$, the contribution of the abiotic component, $ENV_{COMP_i} = inv.logit(e_1 \times env1_i + \cdots e_2 \times env2_i)$ and the contribution of the biotic component, $BIO_{COMP_i} = inv.logit(b_1 \times species1_i + \cdots b_2 \times species2_i$. The logit transform ensures that probabilities close to 0 or 1 are explained equally well by a number of components because the function is flat at extremes. Adapting equations from Beale *et al*. (2014), the proportion of the probability of presence explained by abiotic effects was calculated by:

$$logit(Prop_{ENV_i}) = \frac{|P_i - SE_{COMP_i} - BIO_{COMP_i}|}{|P_i - SE_{COMP_i} - ENV_{COMP_i}| + |P_i - SE_{COMP_i} - BIO_{COMP_i}| + |P_i - BIO_{COMP_i} - ENV_{COMP_i}|}$$

(***Equation* 3**),

And the proportion of the probability of presence explained by biotic factors:

$$logit(Prop_{BIO_i}) = \frac{|P_i - SE_{COMP_i} - ENV_{COMP_i}|}{|P_i - SE_{COMP_i} - BIO_{COMP_i}| + |P_i - SE_{COMP_i} - ENV_{COMP_i}| + |P_i - BIO_{COMP_i} - ENV_{COMP_i}|}$$

(***Equation* 4**).

Outputs from the calculation of $Prop_{BIO_i}$ and $Prop_{ENV_i}$ were plotted spatially by reassigning coordinates to each grid cell. Marginal effect plots for environmental covariates

were produced using original variable values and predicted probability of presence values P_i from INLA models. Marginal effect plots for biotic interactions were produced as boxplots due to the binary input data.

Ecoregional climatic stability was defined as "the proportion of an ecoregion which was predicted to be climatically stable under [future] climate change" (Iwamura *et al.*, 2013). The climatic stability index was calculated by estimating the overlap between present and future climatic envelopes for each ecoregion using results from seven global circulation models (GCMs). It ranges from 0, no overlap between current and future climates, to 1, complete overlap and high robustness to climate change (Watson *et al.*, 2013). Linear regressions were performed to test the relationship between climatic stability and $Prop_{BIO_i}/Prop_{ENV_i}$ for each species.

1.3 RESULTS

Predicted probabilities of presence for four European lagomorph species varied substantially between INLA models built with IUCN geographic range polygons (Fig. 3a) and those built with point occurrence data (Fig. 3b). Although using IUCN data may result in false positives, in this case, point occurrence data resulted in false positive and false negative predictions of occurrence. For example, the range of *L. granatensis* is restricted to the Iberian Peninsula yet models utilising point occurrence predicted areas in Northern Europe to be suitable; erroneously. In addition, *L. europaeus* and *O. cuniculus* are distributed throughout Central Europe extending into Eastern Europe, yet models using point occurrence data predicted distributions skewed to Western Europe. This reflects the sparsity and biased nature of point occurrence data, whilst suggesting that IUCN polygons, at least for this taxon, lead to more realistic species distribution models. However, it is possible that the disparity within species between Fig. 3a and Fig. 3b may in fact be a result of differences in rate of movement within the observed area of occurrence for spatially autocorrelated (Fig. 3a) and uncorrelated distribution patterns (Fig. 3b), with rates likely to be larger for animal species, i.e. lagomorphs, than plants.

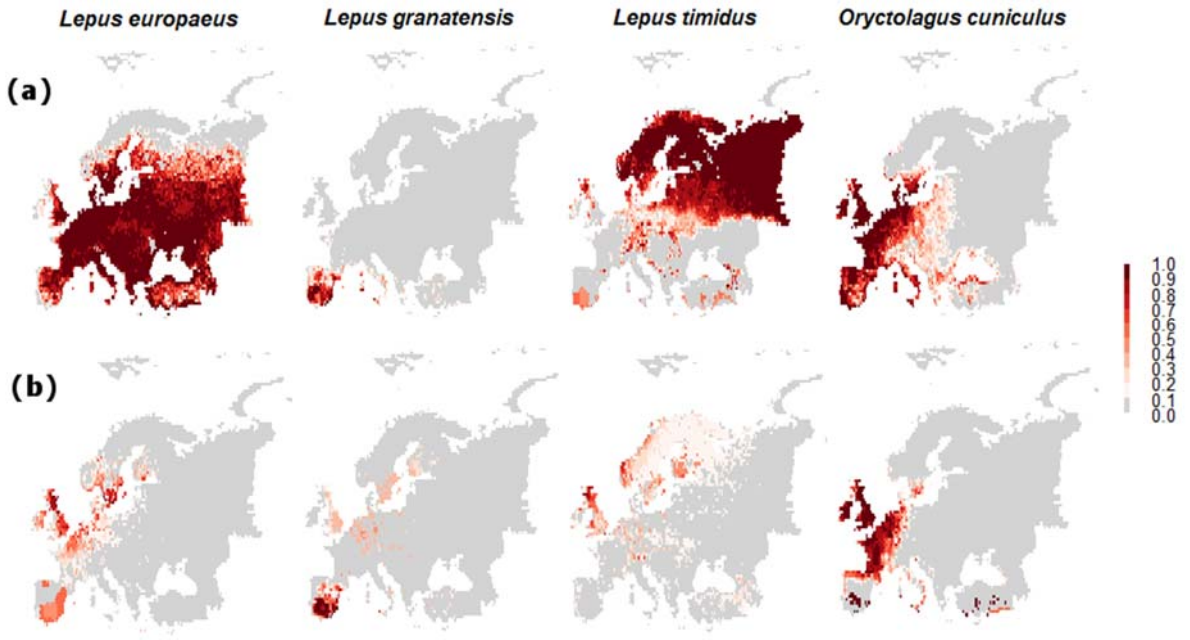


Fig. 3. Predicted probability of presence using **(a)** IUCN geographic range polygons as species input data, and **(b)** using point occurrence data.

Evaluation of the predictive ability of INLA models for four lagomorph species in Europe, as assessed using both WAIC and DIC (which both showed similar results), indicated that those incorporating both abiotic *and* biotic factors were substantially better for all species except *L. timidus* (Table 1). For this species, the abiotic only model had much higher WAIC and DIC values, potentially indicating little influence of biotic factors on its distribution. The difference in WAIC and DIC values for full and abiotic only models of *L. europaeus* and *O. cuniculus* suggested that they differed with a high degree of confidence. Although, the difference was smaller for *L. granatensis*, a ΔWAIC/DIC of ~150 suggests models with abiotic and biotic factors had better predictive power.

Table 1. Predictive ability of INLA models with abiotic and biotic factors compared to models with only abiotic factors. Watanabe-Akaike information criterion (WAIC) and deviance information criterion (DIC) were used to assess predictive ability, where a lower value indicates better predictive power (*) and a greater ΔWAIC/DIC suggests greater differences between the two models.

Species	DIC values			WAIC values		
	Abiotic + Biotic	Abiotic only	ΔDIC	Abiotic + Biotic	Abiotic only	ΔWAIC
Lepus europaeus	2765.06*	3363.69	598.63	2778.80*	3372.87	594.07
Lepus granatensis	420.91*	576.34	155.43	431.32*	581.86	150.54
Lepus timidus	3095.65	2584.00*	-511.65	3095.58	2757.06*	-338.52
Oryctolagus cuniculus	1880.04*	2204.49	324.45	1891.10*	2210.25	319.15

The average influences of abiotic and biotic factors on predicted distributions were largely comparable for all species, with the average proportion of ranges explained by abiotic effects 0.916 and the average proportion explained by biotic factors 0.915, however, the range of *L. granatensis* was explained to a greater extent by abiotic effects (0.907 vs. 0.824) and the range of *O. cuniculus* explained to a greater extent by biotic factors (0.873 vs. 0.925). Though, the proportion of species' presence explained by abiotic (Fig. 4a) and biotic factors (Fig. 4b) varied considerably over space. For example, the range of *L. timidus* was influenced to a greater extent by potential interspecific interactions and to a lesser extent by abiotic factors at its southern range edge, whilst the range of *O. cuniculus* was influenced more by potential interspecific interactions in the western part of its distribution than abiotic factors.

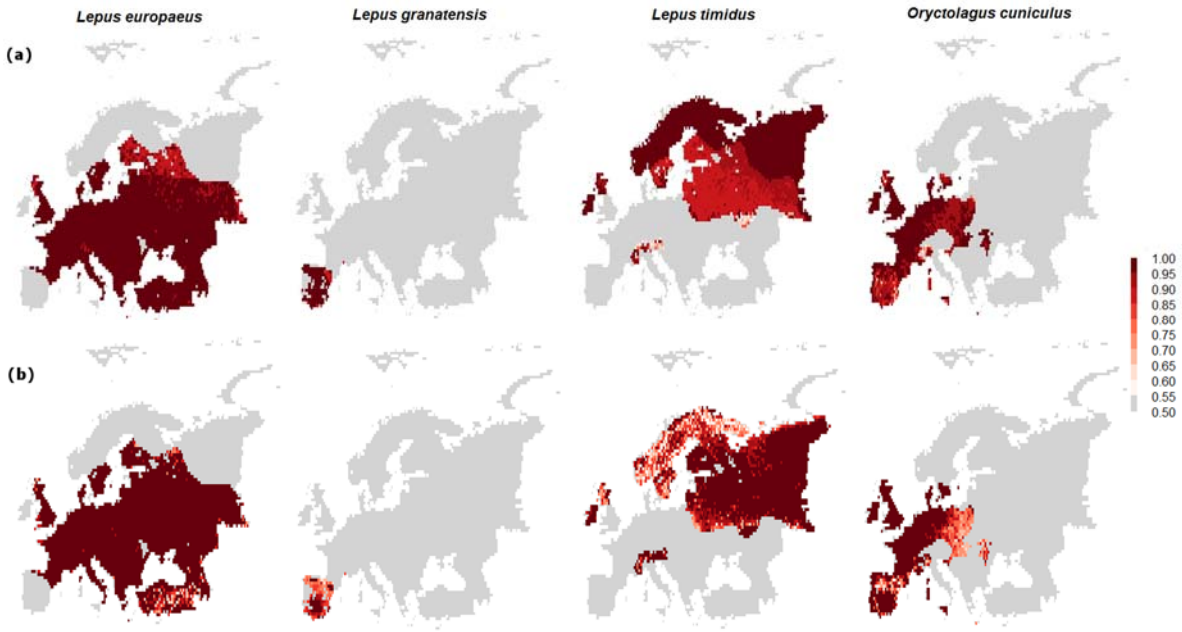


Fig. 4. Proportion of predicted probability of presence explained by **(a)** abiotic factors or **(b)** biotic factors mapped spatially. These maps were clipped to the IUCN geographic range polygons.

Our models suggest the potential for mutualistic as well as competitive interactions within European lagomorph species. *L. granatensis* and *O. cuniculus*, as well as *L. europaeus* and *O. cuniculus*, had higher probabilities of presence where they co-occur with each other, indicating possible mutualism (Fig. 5). Whereas *L. europaeus* had a higher probability of presence in the absence of *L. granatensis* and *L. timidus*, *L. granatensis* and *L. timidus* had higher probabilities in the absence of *L. europaeus*, and *O. cuniculus* had higher predicted probabilities of presence in the absence of *L. timidus* and *vice versa*, indicating possible competition (Fig. 5).

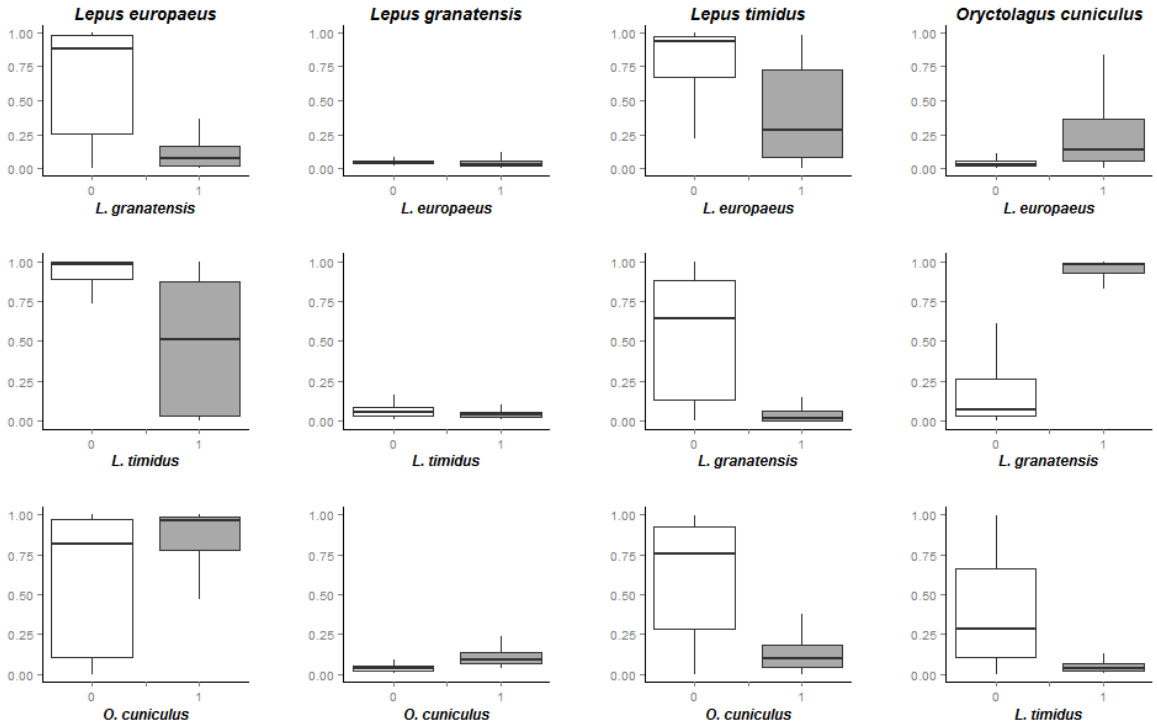
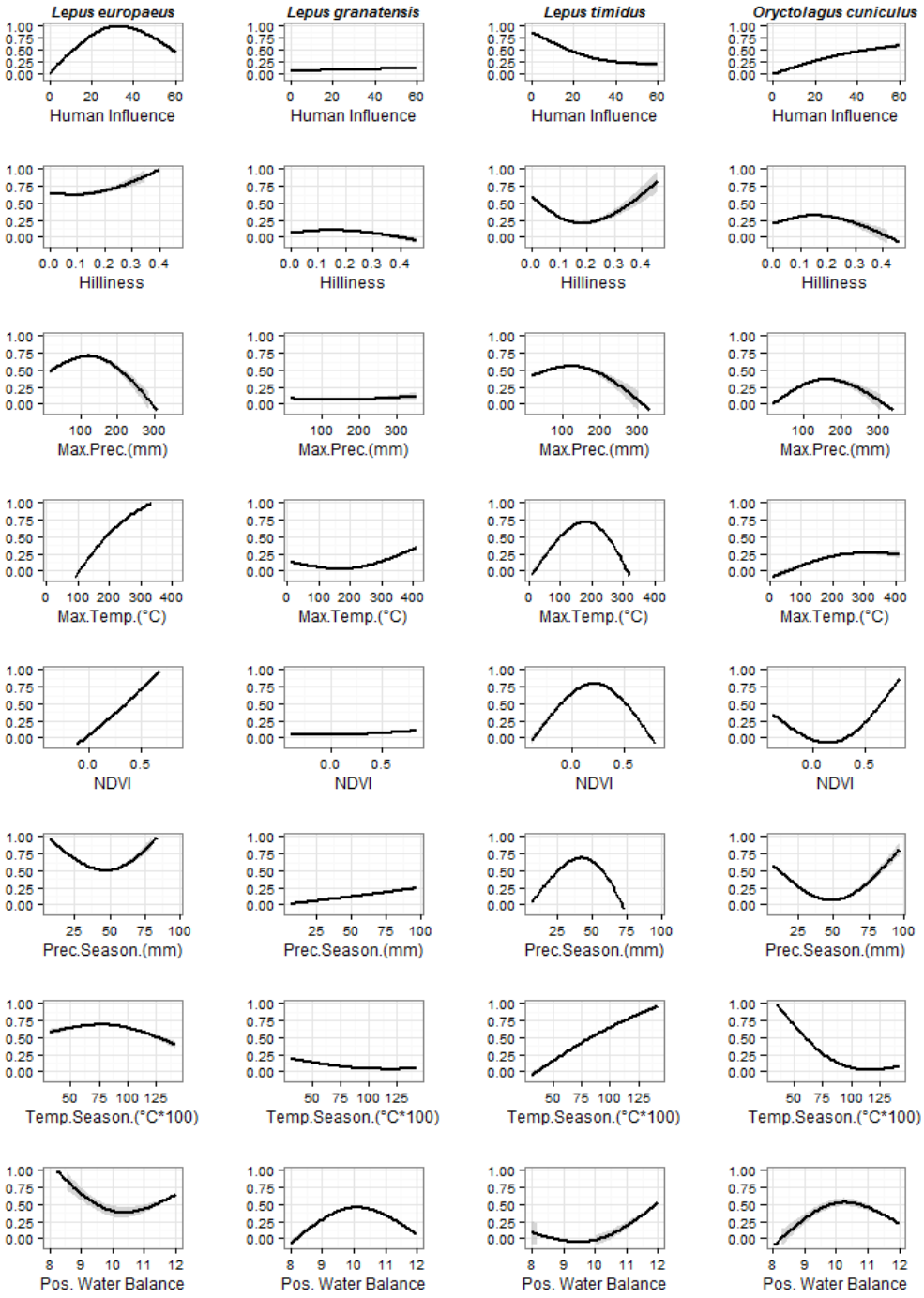


Fig. 5. Marginal effect plots showing the influence of biotic factors, i.e. co-occurring species (where 0 indicates absence and 1 presence), on predicted probabilities of presence for all combinations of species.

In terms of abiotic factors, *L. europaeus* was positively associated with NDVI (high productivity landscapes) and maximum temperature, and more likely to be found in areas of

approximately 100mm maximum precipitation, with little precipitation seasonality (Fig. 6 & Table 2). *L. granatensis* was associated with high maximum temperatures above 20°C. *L. timidus* was positively associated with temperature seasonality, as well as low human influence and extremely hilly areas. *O. cuniculus* was more likely to be present in flat landscapes, at low and high NDVI values i.e. semi-arid regions in Iberia and the Mediterranean, and high productivity landscapes i.e. farmland, and in areas with approximately 150mm maximum precipitation.



347

348 **Fig. 6.** Marginal effect plots showing the relationships between environmental covariates and
349 predicted probabilities of species presence. Grey shaded regions show 95% credible intervals
350 and solid lines the mean estimate.

Table 2. Parameter estimates for abiotic effects on the predicted probabilities of presence of European lagomorph species plus 2.5% and 97.5% credible intervals. Significant parameter estimates, i.e. those with credible intervals that do not cross 0.5, are shown in bold.

Effect		*Lepus europaeus*	*Lepus granatensis*	*Lepus timidus*	*Oryctolagus cuniculus*
Human Influence	Mean	<0.001	0.067	**0.844**	0.117
	97.5%	0.071	0.913	**0.876**	0.152
	2.5%	<0.001	0.003	**0.805**	0.016
Hilliness	Mean	0.502	0.496	**0.513**	**0.509**
	97.5%	0.508	0.500	**0.516**	**0.512**
	2.5%	0.496	0.492	**0.510**	**0.505**
Max. Prec. (mm)	Mean	**0.943**	**0.944**	0.055	**0.941**
	97.5%	**0.969**	**0.999**	0.060	**0.951**
	2.5%	**0.927**	**0.921**	0.047	**0.930**
Max. Temp. (°C)	Mean	**0.704**	**0.787**	<0.001	<0.001
	97.5%	**1.000**	**1.000**	<0.001	<0.001
	2.5%	**0.522**	**0.717**	<0.001	<0.001
NDVI	Mean	**0.648**	0.495	0.474	**0.588**
	97.5%	**0.669**	0.515	0.490	**0.603**
	2.5%	**0.624**	0.474	0.458	**0.573**
Prec. Season. (mm)	Mean	**0.970**	<0.001	0.932	0.989
	97.5%	**0.999**	0.012	0.958	0.999
	2.5%	**0.960**	<0.001	<0.001	0.015
Temp. Season. (°C)	Mean	0.000	0.485	**0.524**	0.503
	97.5%	0.328	1.000	**0.533**	0.516
	2.5%	0.000	0.000	**0.517**	0.493
Positive Water Balance	Mean	0.839	0.364	0.019	0.133
	97.5%	0.985	0.722	0.090	0.468
	2.5%	0.161	0.111	0.010	0.031

The relationship between climatic stability and the proportion of the range influenced by abiotic or biotic factors varied for each species. *L. europaeus* showed a similar response to *O. cuniculus*, and *L. granatensis* a similar response to *L. timidus* despite the geographically distinct distributions (Fig. 7). For *L. europaeus* and *O. cuniculus* areas highly influenced by abiotic factors were positively related to climatic stability (average slope = 0.342), i.e. these

areas are likely to become more climatically stable in the future, whereas areas for those species highly influenced by biotic factors were negatively associated with climatic stability (average slope = -0.142), i.e. these areas are likely to become less climatically stable in the future. For *L. granatensis* and *L. timidus* areas influenced by abiotic factors are predicted to be less stable under future climates (average slope = -0.227) and areas influenced by biotic factors more stable (average slope = 0.240).

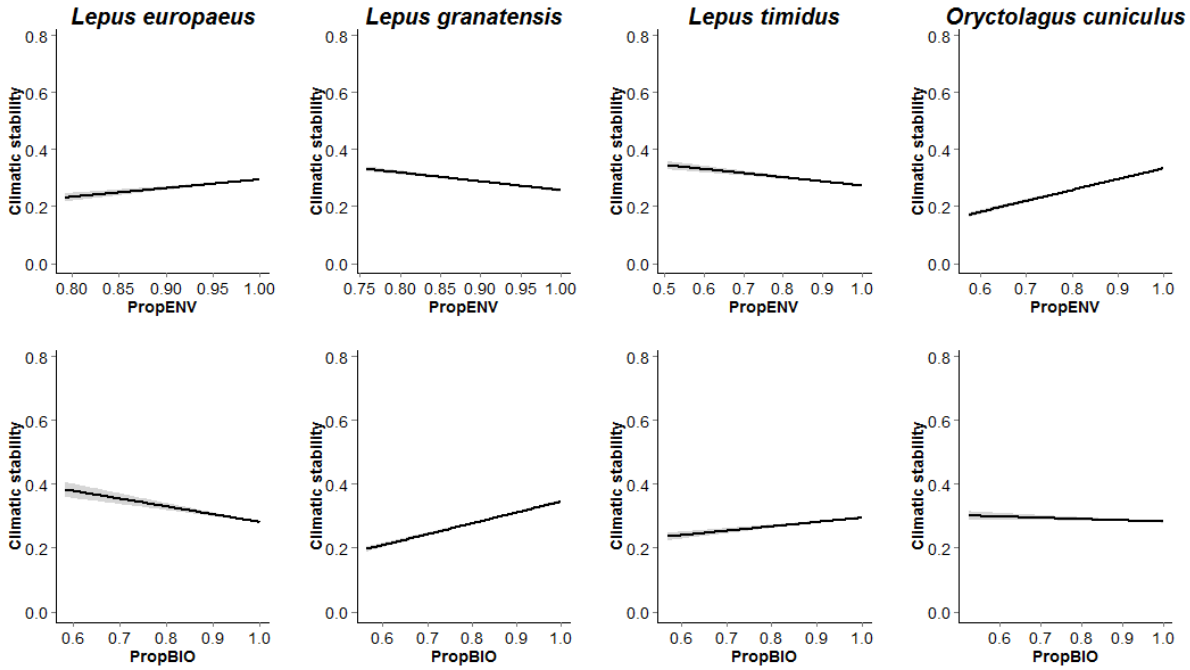


Fig. 7. Relationship between climatic stability and the proportion of predicted probability of presence explained by abiotic (PropENV) or biotic factors (PropBIO) for each species. The grey areas surrounding fitted linear regression lines indicate credible intervals.

1.4 DISCUSSION

Predictive ability was substantially better in models built using abiotic *and* biotic factors, than in models built solely with abiotic variables, for all but one lagomorph, suggesting the potential for interspecific drivers in determining species distribution patterns. Abiotic factors alone are likely to build a better model for the mountain hare, *L. timidus*, due to its distribution being primarily driven by cooler temperatures at high latitudes and elevations. For the majority of lagomorph species in Europe, abiotic factors and interspecific interaction effects appeared *equally* important in determining their ranges. Our models suggest directional associations consistent with field-based observations reported in the literature between *L. granatensis* and *L. europaeus* (Gortazar *et al*., 2007), *L. europaeus* and *L. timidus* (Thulin, 2003; Reid & Montgomery, 2007; Reid, 2011; Caravaggi *et al.*, 2014; 2016) and *O. cuniculus* and *L. timidus* (though this latter relationship has not been empirically documented their ranges show significant overlap and biotic interactions are both possible and probable, particularly at the southerly range edge of *L. timidus*). In addition, a potential mutualistic interaction is suggested between *O. cuniculus* and *L. granatensis*, which to the best of our knowledge has not yet been reported. However, these correlative analyses may only identify associations which cannot be used to attribute causation; underlying mechanisms, for example, the strength and direction of interspecific competition, can only be established using field observations and/or hypothesis testing experiments (Beale *et al*., 2014). Furthermore, these results do not consider the influence of additional biotic factors, for example pikas (*Ochotona roylei*) and birds (*Pnoepyge albiventer*) are known to exhibit mutualism in some locations (Khanal, 2007), and also do not consider that possible mutualistic relationships could be due to competitive interactions of species in the same niche, not reaching equilibrium.

The influence of biotic factors was spatially heterogeneous. For example, it is predicted that biotic factors disproportionately influenced the southern range of *L. timidus*, a predominately

high latitude species which overlaps with its competitor *L. europaeus*, a predominately temperate climate, lowland farmland species at its southern range margin (Thulin, 2003). The western regions of the range of *O. cuniculus* were also suggested to be disproportionately influenced by biotic factors presumably where it overlaps with other lagomorphs.

Environmental covariates identified as important to the predicted presence of European lagomorphs in our models were largely supported in the literature. *O. cuniculus* prefers flat landscapes (Fa *et* al., 1999; Tapia *et al*., 2010), and *L. europaeus* and *O. cuniculus* occupy landscapes of differing productivity and vegetation (Smith & Boyer, 2008; Smith & Johnston, 2008). In addition, *L. europaeus* and *O. cuniculus* prefer less seasonal and lower levels of precipitation (Rödel & Dekker, 2012), whereas *L. granatensis* is only found in areas of relatively high temperatures (Alves & Rocha, 2003). *L. timidus* occupies regions with high winter snowfall (Angerbjorn & Flux, 1995), and, therefore, tolerates large seasonal variations in temperature. Also, *L. timidus* is known to occupy high elevations, especially in the Alps, and inhabits areas of low human influence (Thulin, 2003).

Projecting species distributions under future climate scenarios is inherently challenging, especially with the inclusion of biotic factors which may introduce the potential for multicollinearity (Kissling *et al*., 2012). Associating proxies for climate change with the results of spatial hierarchical models can provide insight into the potential impacts of future change (e.g. Watson *et al*., 2013). We show that areas of *O. cuniculus* and *L. europaeus* ranges influenced by biotic factors will be less robust to future changes in climate, whereas areas of *L. granatensis* and *L. timidus* ranges highly influenced by the environment will be less robust to future climate. The narrow climatic conditions occupied by the latter two species and the uniqueness of these conditions within Europe, i.e. high temperatures in Iberia (Alves & Rocha, 2003) and cold temperatures in Northern Europe (Thulin, 2003), probably explains why the areas influenced by abiotic factors are less robust – these particularly arid and arctic ecoregions

are highly susceptible to even small changes in climate (Beaumont *et al*., 2011), and the species found here, which are adapted to these narrow conditions, may therefore struggle to cope with small changes to temperature or precipitation. The widespread lagomorphs, *L. europaeus* and *O. cuniculus*, on the other hand, are predicted to be less robust to future climatic changes in areas determined by biotic factors. This may lead to substantial changes in future distributions, given that these species co-occur with most other lagomorph species in Europe, and may lead to altered interactions, for example one of the more range-restricted species may outcompete one of the widespread species in the future, potentially changing the direction of interactions. The effects on range-restricted species, e.g. *L. castroviejoi* and *L. corsicanus*, are still unknown, but we expect them to be particularly sensitive to projected changes due to the restricted climatic envelopes that they occupy. On the basis of these results we suggest that the management of *L. timidus* and *L. granatensis* in the future is directed towards habitat conservation in areas of climate particularly favoured by these species, whereas for *L. europaeus* and *O. cuniculus* conservation strategies should be directed at areas where these species overlap with co-occurring lagomorphs, for example conservation of *O. cuniculus* in Iberia, France and the UK, and *L. europaeus* in the UK, Central Europe and the more southerly regions of its overlap with *L. timidus*.

Although our species-specific results are consistent with environmental and interspecific effects reported in the literature, given unbiased and adequate point occurrences with reliable data on observer effort, models could be improved (see Royle *et* al., 2007; Beale *et* al., 2014). Here, we concentrated on determining interspecific interaction effects at the same trophic level with congeneric species (for simplicity), but future improvements might also include top-down (predator-prey), bottom-up (plant-grazer) and host-parasite interactions as these are equally, if not more likely, to have strong impacts on species distributions (Kissling *et al*., 2012).

1.4.1 Conclusions

The approach demonstrated here allowed biotic factors to be included in modelling species distributions, without the need for *a priori* knowledge on species interactions. SDMs that do not explicitly include biotic, as well as abiotic factors are missing a large source of variation that appears, for European lagomorphs, equally important in determining species distributions. Even though the results are limited to European lagomorphs and the approach might not be relevant for some taxa, for example plants, it is likely that it can be applied to a wide range of species given adequate species occurrence and environmental data. In addition, through the production of accurate species distribution models, built with abiotic and biotic factors, effective and appropriate conservation measures can be suggested. However, these approaches must be extended to model rare, and/or highly range-restricted species which are likely to be particularly vulnerable in the future, yet are seldom included due to sparse data and poor model performance. The evaluation of risk to future conservation status of a species by extrapolation under future climate scenarios of SDMs that do not include biotic factors is likely to result in biased and unrealistic results (except in those cases where the effect of biotic interactions is small, which we suspect likely to be the case for a minority of species only). Species distributions largely defined by biotic interactions cannot be extrapolated without simultaneously extrapolating the distribution of all potentially interacting species, including top-down and bottom-up processes. Even if the statistical challenges of creating such complex models are overcome, they are unlikely to account for changes in the strength and direction of biotic interactions in the future because species range determinants are likely to have multivariate influences. Thus, the development of Joint SDM approaches and network analysis to quantify the strength of biotic interactions at all trophic levels accounting for spatial autocorrelation whilst minimising computational power is needed.

ACKNOWLEDGEMENTS

This project was funded by *Quercus*, Northern Ireland's Centre for Biodiversity and Conservation Science supported by the School of Biological Sciences, Queen's University Belfast (QUB). We would like to thank Jack Lennon and Hannah White for initial discussion of the manuscript.

REFERENCES

Acevedo, P., Jimenez-Valverde, A., Melo-Ferreira, J., Real, R. & Alves, P.C. (2012) Parapatric species and the implications for climate change studies: a case study on hares in Europe. *Global Change Biology,* **18**, 1509-1519.

Allouche, O., Tsoar, A. & Kadmon, R. (2006) Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, **43**, 1223-1232.

Alves, P.C. & Rocha, A. (2003) Environmental factors have little influence on the reproductive activity of the Iberian hare (*Lepus granatensis*). *Wildlife Research*, **30**, 639-647.

Amarasekare, P. (2003) Competitive coexistence in spatially structured environments: a synthesis. *Ecology Letters,* **6**, 1109-1122.

Angerbjorn, A. & Flux, J.E.C. (1995) *Lepus timidus*. *Mammalian Species,* **495**, 1-11.

Araujo, M.B. & Luoto, M. (2007) The importance of biotic interactions for modelling species distributions under climate change. *Global Ecology & Biogeography*, **16**, 743–753.

Austin, M.P. (2002) Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecological Modelling*, **157**, 101-118.

Beaumont, L.J., Pitman, A., Perkins, S., Zimmerman, N.E., Yoccoz, N.G. & Thuiller, W. (2011) Impacts of climate change on the world's most exceptional ecoregions. *Proceedings of the National Academy of Sciences USA*, **108**, 2306-2311.

496 Beguin, J., Martino, S., Rue, H. & Cumming, S.G. (2012) Hierarchical analysis of spatially
497 autocorrelated ecological data using integrated nested Laplace approximation. *Methods in*
498 *Ecology & Evolution*, **3**, 921–929.

499 Beale, C.M., Lennon, J.J., Yearsley, J.M., Brewer, M.J. & Elston, D.A. (2010) Regression
500 analysis of spatial data. *Ecology Letters,* **13**, 246-264.

501 Beale, C.M., Brewer, M.J. & Lennon, J.J. (2014) A new statistical framework for the
502 quantification of covariate associations with species distributions. *Methods in Ecology &*
503 *Evolution*, **5**, 421–432.

504 Bull, C.M. (1991) Ecology of parapatric distributions. *Annual Review of Ecology &*
505 *Systematics*, **22**, 19-36.

506 Caravaggi, A., Montgomery, W.I. & Reid, N. (2014) Range expansion and comparative habitat
507 use of insular, congeneric lagomorphs: invasive European hares *Lepus europaeus* and
508 endemic Irish hares *Lepus timidus hibernicus*. *Biological Invasions*, **17**, 687-698.

509 Caravaggi, A., Zaccaroni, M., Riga, F., Schai-Braun, S.C., Dick, J.T.A., Montgomery, W.I. &
510 Reid, N. (2016) An invasive-native mammalian species replacement process captured by
511 camera trap survey random encounter models. *Remote Sensing in Ecology and*
512 *Conservation*, **2**, 45-58.

513 Cornell, H.V. & Lawton, J.H. (1992) Species interactions, local and regional processes and
514 limits to the richness of ecological communities: a theoretical perspective. *Journal of*
515 *Animal Ecology*, **61**, 1-12.

516 Elith, J., Graham, C.H., Anderson, R.P. *et al*. (2006) Novel methods improve prediction of
517 species' distributions from occurrence data. *Ecography*, **29**, 129–151.

518 Elith, J. & Leathwick, J.R. 2009. Species distribution models: ecological explanation and
519 prediction across space and time. *Annual Review of Ecology, Evolution & Systematics*, **40**,
520 677-697.

521 Fa, J.E., Sharples, C.M. & Bell, D.J. (1999) Habitat correlates of European rabbit (*Oryctolagus*
522 *cuniculus*) distribution after the spread of RVHD in Cadiz Province, Spain. *Journal of*
523 *Zoology*, **249**, 83-96.

524 Flux, J.E.C. (2008) A review of competition between rabbits (*Oryctolagus cuniculus*) and hares
525 (*Lepus europaeus*). In: Alves, P.C., Ferrand, N. & Hackländer, K. (eds) *Lagomorph*
526 *Biology: Evolution, Ecology and Conservation,* 241-249. Springer-Verlag Berlin
527 Heidelberg.

528 Gelman, A., Hwang, J. & Vehtari, A. (2014) Understanding predictive information criteria for
529 Bayesian models. *Statistical Computing*, **24**, 997-1016.

530 Gilman, S.E., Urban, M.C., Tewksbury, J., Gilchrist, G.W. & Holt, R.D. (2010) A framework
531 for community interactions under climate change. *Trends in Ecology & Evolution*, **25**, 325-
532 331.

533 Gortazar, C., Millán, J., Acevedo, P., Escudero, M.A., Marco, J. & Fernández de Luco, D.
534 (2007) A large-scale survey of brown hare *Lepus europaeus* and Iberian hare *L. granatensis*
535 populations at the limit of their ranges. *Wildlife Biology*, **13**, 244-250.

536 Graham, C.H. & Hijmans, R.J. (2006) A comparison of methods for mapping species ranges
537 and species richness. *Global Ecology & Biogeography*, **15**, 578–587.

538 Gutiérrez, E.E., Boria, R.A. & Anderson, R.P. (2014) Can biotic interactions cause allopatry?
539 Niche models, competition and distributions of South American mouse opossums.
540 *Ecography*, **37**, 741-753.

541 Harris, D.J. (2014) Generating realistic assemblages with a Joint Species Distribution Model.
542 *Methods in Ecology & Evolution*, **6**, 465-473.

543 Hof, A.R., Jansson, R. & Nilsson, C. (2012) How biotic interactions may alter future
544 predictions of species distributions: future threats to the persistence of the arctic fox in
545 Fennoscandia. *Diversity & Distributions*, **18**, 554-562.

546 Illian, J. (2013) Fitting complex models in INLA – developments and extensions. Accessed on
547 24/03/2015 at http://www.bioss.ac.uk/rsse/2013/September2013slides-Illian.pdf.

548 Iwamura, T., Wilson, K.A., Venter, O. & Possingham, H.P. (2010) A Climatic Stability
549 Approach to Prioritizing Global Conservation Investments. *PLOS ONE*, **5**, e15103.

550 Iwamura, T., Guisan, A., Wilson, K.A. & Possingham, H.P. (2013) How robust are global
551 conservation priorities to climate change? *Global Environmental Change*, **23**, 1277-1284.

552 Jablonski, D. 2008. Biotic interactions and macroevolution: extensions and mismatches across
553 scales and levels. *Evolution*, **62**, 715-739.

554 Khanal, B. 2007. New Report on the Symbiotic Relation of *Ochotona roylei* (Lagomorpha:
555 Ochotonidae) and Scaly Breasted Wren Babbler (*Pnoepyge albiventer*) at Ganesh Himalaya
556 Area of Central Nepal. *Our Nature*, **5**, 37-40.

557 Kissling, W.D, Dormann, C.F, Groeneveld, J. *et al*. (2012) Towards novel approaches to
558 modelling biotic interactions in multispecies assemblages at large spatial extents. *Journal*
559 *of Biogeography*, **39**, 2163–2178.

560 Kremen, C., Cameron, A., Moilanen, A. *et al*. (2008) Aligning conservation priorities across
561 taxa in Madagascar with high-resolution planning tools. *Science*, **320**, 222–226.

562 Lawler, J.J., Shafer, S.L., White, D., Kareiva, P., Maurer, E.P., Blaustein, A.R. & Bartlein, P.J.
563 (2009) Projected climate-induced faunal change in the Western Hemisphere. *Ecology*, **90**,
564 588-597.

565 Leach, K., Montgomery, W.I. & Reid, N. (2015a) Biogeography, macroecology and species
566 traits mediate competitive interactions in the Order Lagomorpha. *Mammal Review*, **45**, 88-
567 102.

568 Leach, K., Kelly, R., Cameron, A., Montgomery, W.I. & Reid, N. (2015b) Expertly validated
569 models and phylogenetically-controlled analysis suggests responses to climate change are
570 related to species traits in the Order Lagomorpha. *PLOS ONE*, **10**, e0122267.

571 Legendre, P. (1993) Spatial autocorrelation: trouble or new paradigm? *Ecology*, **74**, 1659-1673.

572 Lennon, J.J. (2000) Red-shifts and red herrings in geographical ecology. *Ecography*, **23**, 101-
573 113.

574 Maestre, Z.A., Bowker, M.A., Escolar, C. *et al*. (2010) Do biotic interactions modulate
575 ecosystem functioning along stress gradients? Insights from semi-arid plant and biological
576 soil crust communities. *Philosophical Transactions of the Royal Society London Series B*,
577 **365**, 2057-2070.

578 Midgley, G.F., Davies, I.F., Albert, C.H. *et al*. (2010) BioMove: an integrated platform
579 simulating the dynamic response of species to environmental change. *Ecography*, **33**, 612-
580 616.

581 Mod, H.K., le Roux, P.C., Guisan, A. & Luoto, M. (2015) Biotic interactions boost spatial
582 models of species richness. *Ecography*, **38**, 913-921.

583 Murray, K.A., Rosauer, D., McCallum, H. & Skerratt, L.F. (2011) Integrating species traits
584 with extrinsic threats: closing the gap between predicting and preventing declines.
585 *Proceedings of the Royal Society of London B: Biological Sciences*, **278**, 1515-1523.

586 Newton-Cross, G., White, P.C.L. & Harris, S. (2007) Modelling the distribution of badgers
587 *Meles meles*: comparing predictions from field-based and remotely derived habitat data.
588 *Mammal Review*, **37**, 54–70.

589 Pearson, R.G. & Dawson, T.P. (2003) Predicting the impacts of climate change on the
590 distribution of species: are bioclimate envelope models useful? *Global Ecology &*
591 *Biogeography*, **12**, 361-371.

592 Pellissier, L., Anne Bråthen, K., Pottier, J., Randin, C. F., Vittoz, P., Dubuis, A., Yoccoz, N.
593 G., Alm, T., Zimmermann, N. E. & Guisan, A. (2010) Species distribution models reveal
594 apparent competitive and facilitative effects of a dominant species on the distribution of
595 tundra plants. *Ecography*, **33**, 1004–1014.

596 Plummer, M. (2008) Penalized loss functions for Bayesian model comparison. *Biostatistics*, **9**,
597 523-539.

598 Pollock, L.J., Tingley, R., Morris, W.K., Golding, N., O'Hara, R.B., Parris, K.M., Vesk, P.A.
599 & McCarthy, M.A. (2014) Understanding co-occurrence by modelling species
600 simultaneously with a Joint Species Distribution Model (JSDM). *Methods in Ecology &*
601 *Evolution*, **5**, 397–406.

602 Reid, N. (2011) European hare (*Lepus europaeus*) invasion ecology: implication for the
603 conservation of the endemic Irish hare (*Lepus timidus hibernicus*). *Biological Invasions*, **13**,
604 559-569.

605 Reid, N. & Montgomery, W.I. (2007) Is naturalisation of the brown hare in Ireland a threat to
606 the endemic Irish hare? *Biology and Environment: Proceedings of the Royal Irish Academy*,
607 **107B**, 129-138.

608 Rödel, H.G. & Dekker, J.J.A. (2012) Influence of weather factors on population dynamics of
609 two lagomorph species, based on hunting bag records. *European Journal of Wildlife*
610 *Research*, **58**, 923-932.

611 Royle, J.A., Kéry, M., Gautier, R. & Schmid, H. (2007) Hierarchical Spatial Models of
612 Abundance and Occurrence from Imperfect Survey Data. *Ecological Monographs*, **77**, 465–
613 481.

614 Rue, H., Martino, S. & Chopin N. (2009) Approximate Bayesian Inference for Latent Gaussian
615 Models Using Integrated Nested Laplace Approximations. *Journal of the Royal Statistical*
616 *Society: Series B*, **71**, 319–392.

617 Smith, A.T. & Boyer, A.F. (2008). *Oryctolagus cuniculus*. The IUCN Red List of Threatened
618 Species. Version 2014.3.

619 Smith, A.T. & Johnston, C.H. (2008). *Lepus europaeus*. The IUCN Red List of Threatened
620 Species. Version 2014.3.

621 Soberon, J. & Nakamura, M. (2009) Niches and distributional areas: concepts, methods and
622 assumptions. *Proceedings of the National Academy of Sciences USA*, **106**, 19644-19650.

623 Spiegelhalter, D.J., Best, N.G., Carlin, B.P. & Van Der Linde, A. (2002) Bayesian measures of
624 complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical*
625 *Methodology)*, **64**, 583-639.

626 Tapia, L., Dominguez, J. & Rodriguez, L. (2010) Modelling habitat use by Iberian hare *Lepus*
627 *granatensis* and European wild rabbit *Oryctolagus cuniculus* in a mountainous area in
628 north-western Spain. *Acta Theriologica*, **55**, 73-79.

629 Thuiller, W., Araujo, M.B. & Lavorel, S. (2003) Generalized models vs. classification tree
630 analysis: predicting spatial distributions of plant species at different scales. *Journal of*
631 *Vegetation Science*, **14**: 669-680.

632 Thulin, C-G. (2003) The distribution of mountain hares *Lepus timidus* in Europe: a challenge
633 from brown hares *L. europaeus*? *Mammal Review*, **33**, 29-42.

634 Visconti, P., Bakkenes, M., Baisero, D. *et al*. (2015) Projecting global biodiversity indicators
635 under future development scenarios. *Conservation Letters*, doi: 10.1111/conl.12159.

636 Watanabe, S. (2010) Asymptotic equivalence of Bayes cross validation and Widely Applicable
637 Information Criterion in singular learning theory. *Journal of Machine Learning Research*,
638 **11**, 3571-3594.

639 Watson, J.E.M., Iwamura, T. & Butt, N. (2013) Mapping vulnerability and conservation
640 adaptation strategies under climate change. *Nature Climate Change*, **3**, 989-994.

641 WCS (Wildlife Conservation Society) CIESIN (Centre for International Earth Science
642 Information Network). (2005) *Last of the Wild Project, Version 2, 2005: Global Human*
643 *Influence Index Dataset.* Pallisades, NY: NASA Socioeconomic Data and Applications
644 Centre. http://doi.org/10.7927/H46W980H.

645 Whittaker, R.J. (1975) *Communities and Ecosystems*. MacMillan, New York.

646 Wiens, J.J. (2011) The niche, biogeography and species interactions. *Philosophical*
647 *Transactions of the Royal Society London Series B*, **366**, 2336-2350.

648 Wisz, M.S., Pottier, J., Kissling, W.D. *et al*. (2012) The role of biotic interactions in shaping
649 distributions and realised assemblages of species: implications for species distribution
650 modelling. *Biological Reviews*, **88**, 15-30.